

A Non-Commutative Formula for the Isotropic Magneto-Electric Response

Bryan Leung

Department of Physics & Astronomy, Rutgers University, Piscataway, NJ 08854,
USA
bryleung@physics.rutgers.edu

Emil Prodan

Department of Physics, Yeshiva University, New York, NY 10016, USA
prodan@yu.edu

Abstract. A non-commutative formula for the isotropic magneto-electric response of disordered insulators under magnetic fields is derived using the methods of non-commutative geometry. Our result follows from an explicit evaluation of the Ito derivative with respect to the magnetic field of the non-commutative formula for the electric polarization reported in Ref. [1]. The quantization, topological invariance and connection to a second Chern number of the magneto-electric response are discussed in the context of 3-dimensional, disordered, time-reversal or inversion symmetric topological insulators.

PACS numbers: 02.30.Sa, 75.85.+t, 71.23.-k

1. Introduction

The magneto-electric effect in insulating materials consists in the appearance of a finite, measurable electric bulk polarization \mathbf{P} when a sample is subjected to an external magnetic field \mathbf{B} , and in the appearance of a finite, measurable bulk magnetization \mathbf{M} when the sample is subjected to an external electric field \mathbf{E} . The magneto-electric effect is quantified by the magneto-electric response tensor:

$$\alpha_{ij} = \frac{\partial P_i}{\partial B_j} = \frac{\partial M_j}{\partial E_i}, \quad (1)$$

where the derivatives are not necessarily taken at zero \mathbf{B} or \mathbf{E} fields. The effect has been observed in a variety of materials and its technological applications can be tremendous [2, 3].

Originally, the effect was sought in materials with broken time-reversal and inversion symmetries, but recently it was shown that certain classes of time-reversal or inversion invariant systems can display large magneto-electric responses [4, 5, 6] due to topological characteristics of their electronic structures. Such materials, which are now routinely synthesized and characterized in laboratories [7, 8], are called 3-dimensional topological insulators. It emerged in the recent years that the 3-dimensional topological insulators can be classified by their magneto-electric response. For example, for perfectly periodic, time-reversal invariant insulators, the magneto-electric response induces a \mathbb{Z}_2 topological classification [4] and the topological part of the magneto-electric tensor, which can serve as the topological invariant for this classification, was shown [9] to equal the previously introduced [10, 11, 12] \mathbb{Z}_2 invariants. This was an important development in the field of topological insulators because it provided a measurable bulk effect which sets apart the trivial and topological 3-dimensional time-reversal invariant insulators [13].

The magneto-electric response has ionic, spin and orbital contributions. Here we discuss only the orbital component, which, at its turn, can be decomposed into isotropic and traceless parts:

$$\alpha_{ij} = \delta_{ij} \frac{1}{3} \sum_{k=1}^3 \alpha_{kk} + \tilde{\alpha}_{ij}. \quad (2)$$

Both components are experimentally measurable quantities but we will focus exclusively on the isotropic component:

$$\alpha = \frac{1}{3} \sum_{j=1}^3 \alpha_{jj}, \quad (3)$$

for it contains the topological response. For perfectly periodic, time-reversal or inversion invariant crystals at zero external fields, α was derived and explicitly computed in Refs. [4, 14, 6] for several lattice models of time-reversal or inversion symmetric insulators, showing that indeed α can take non-trivial quantized values. For perfectly periodic systems without time-reversal and inversion symmetries, Refs. [15, 16] showed that α has additional contributions besides the topological component.

Our goal here is to extend the above mentioned works to the case when disorder and magnetic fields are present, maintaining the assumption of a gap in the energy spectrum at the Fermi level. The main challenges arising in this more general setup are the breaking of the translational symmetry and the tedious linear response calculus

with respect to the magnetic field. Both challenges are overcome with the use of non-commutative geometric methods [17]. The starting point for our work is Ref. [1] where non-commutative formulas for \mathbf{P} and \mathbf{M} have been reported. Given these results, we could have proceeded along two routes, one relying on the variation of \mathbf{P} with \mathbf{B} and the other one relying on the variation of \mathbf{M} with \mathbf{E} . We chose the first route because computing the variation of \mathbf{P} with \mathbf{B} fits better into the non-commutative geometry of the Brillouin torus [18], which provides a set of explicit rules of calculus, including rules for how to compute the derivatives with respect to the magnetic field (see Sections 3 and 4). These rules of calculus alone enabled us to complete the task at hand.

For disordered time-reversal or inversion symmetric insulators ($\mathbf{B} = 0$), under the spectral gap assumption, we are able to formulate the following statement (similar to the existing one for perfectly periodic crystals [4, 5, 6]): The total change $\Delta\alpha[\gamma]$ of α , induced by an adiabatic deformation along a path γ in the model's parameter space, is equal to half of the second Chern number over the manifold $(\gamma - \Sigma\gamma) \times \text{Brillouin torus}$ (when the units are chosen such that $e = h = 1$). Here, Σ is either the time-reversal operation or the inversion operation and the path γ starts and ends with a time-reversal or inversion symmetric system. Since $\Delta\alpha[\gamma]$ can change only by an integer number if the path γ is re-routed, it follows that the integer or half-integer character of $\Delta\alpha[\gamma]$ is independent of the deformation path and this extends the well-known \mathbb{Z}_2 topological classification to the disordered case. We would like to mention that the methods of non-commutative geometry enabled us to derive this result directly in the thermodynamic limit and to avoid the use of arguments based on twisted boundary conditions on large but finite volumes. The latter arguments are hard to control in the thermodynamic limit and they are problematic when strong disorder is considered.

The explicit formula for α (Eq. 47) and its topological properties discussed in the present work is a small step in a program that aims at elucidating what happens when strong disorder is present. By strong disorder we mean situations when the Fermi level is embedded in a dense localized spectrum rather than being positioned in a spectral gap. The crucial question to answer is if the isotropic magneto-electric response remains quantized and invariant under the assumption of a mobility gap (as opposed to a spectral gap) at the Fermi level. A rigorous answer to this question will finally resolve the problem of existence (or lack) of robust extended states in 3-dimensional topological insulators.

2. Disordered lattice-models in the presence of magnetic fields

The analysis reported in this work applies to generic disordered lattice-models in 3-dimensions. The Hilbert space for such a model is spanned by vectors of the form $|\mathbf{n}, \alpha\rangle$, where $\mathbf{n} \in \mathbb{Z}^3$ is a node of the lattice and $\alpha = 1, \dots, D$ labels the orbitals associated with that node. The scalar product $\langle \cdot, \cdot \rangle$ is by definition such that $\langle \mathbf{n}, \alpha | \mathbf{m}, \beta \rangle = \delta_{\mathbf{n}\mathbf{m}} \delta_{\alpha\beta}$. The generic lattice-Hamiltonians, in the absence of a magnetic field and disorder, take the form:

$$H_0 = \sum_{\mathbf{n}, \alpha} \sum_{\mathbf{m}, \beta} t_{\mathbf{n}-\mathbf{m}}^{\alpha\beta} |\mathbf{n}, \alpha\rangle \langle \mathbf{m}, \beta|, \quad (4)$$

Accurate, material-specific lattice-models can be generated from *ab-initio* computations using Wannier function representations or simply by fitting the measured band structure [19]. In almost all cases, the lattice-Hamiltonians involve only a finite number of hoppings between one unit cell and its neighbors, which is something we will

assume in the followings.

In the presence of disorder and a uniform magnetic field \mathbf{B} , which will be parametrized as a 3×3 antisymmetric tensor $\hat{\mathbf{B}}$:

$$\hat{\mathbf{B}} = \begin{pmatrix} 0 & B_3 & -B_2 \\ -B_3 & 0 & B_1 \\ B_2 & -B_1 & 0 \end{pmatrix}, \quad (5)$$

the lattice-Hamiltonian becomes:

$$H_\omega = \sum_{\mathbf{n}, \alpha} \sum_{\mathbf{m}, \beta} e^{i\pi(\mathbf{n}, \hat{\mathbf{B}}\mathbf{m})} t_{\mathbf{n}-\mathbf{m}}^{\alpha\beta}(\omega) |\mathbf{n}, \alpha\rangle \langle \mathbf{m}, \beta|, \quad (6)$$

where the phase factor $e^{i\pi(\mathbf{n}, \hat{\mathbf{B}}\mathbf{m})}$ encodes the effect of the magnetic field through the Peierls substitution [20]. Above, (\cdot, \cdot) denotes the 3-dimensional Euclidean scalar product and $i = \sqrt{-1}$. The hopping amplitudes now have a random component, with a typical form:

$$t_{\mathbf{n}-\mathbf{m}}^{\alpha\beta}(\omega) = t_{\mathbf{n}-\mathbf{m}}^{\alpha\beta} + \lambda \omega_{\mathbf{m}, \mathbf{n}}, \quad (7)$$

where $\omega_{\mathbf{m}, \mathbf{n}}$ are independent random variables uniformly distributed in the interval $[-\frac{1}{2}, \frac{1}{2}]$. The collection of all random variables $\omega = \{\omega_{\mathbf{m}, \mathbf{n}}\}_{\mathbf{m}, \mathbf{n}}$ can be viewed as a point in an infinite dimensional configuration space Ω , which can be equipped with a probability measure:

$$dP(\omega) = \prod_{\mathbf{m}, \mathbf{n}} d\omega_{\mathbf{m}, \mathbf{n}}. \quad (8)$$

Furthermore, the discrete \mathbb{Z}^3 additive group has a natural action on Ω given by:

$$(\mathbf{t}_{\mathbf{a}}\omega)_{\mathbf{m}, \mathbf{n}} = \omega_{\mathbf{m}+\mathbf{a}, \mathbf{n}+\mathbf{a}}, \quad (9)$$

which acts ergodically on Ω and leaves $dP(\omega)$ invariant. If $U_{\mathbf{a}}$ denotes the magnetic translation by \mathbf{a} :

$$U_{\mathbf{a}} |\mathbf{n}, \alpha\rangle = e^{i\pi(\mathbf{a}, \hat{\mathbf{B}}\mathbf{n})} |\mathbf{n} - \mathbf{a}, \alpha\rangle, \quad (10)$$

then $\{H_\omega\}_{\omega \in \Omega}$ forms a covariant family of operators in the sense that:

$$U_{\mathbf{a}} H_\omega U_{\mathbf{a}}^{-1} = H_{\mathbf{t}_{\mathbf{a}}\omega}. \quad (11)$$

In the above settings, the quadruple $(\Omega, dP(\omega), \{\mathbf{t}_{\mathbf{a}}\}_{\mathbf{a} \in \mathbb{Z}^3}, \{H_\omega\}_{\omega \in \Omega})$ determines a homogenous system [18], for which the intensive thermodynamic variables are, with probability one, independent of the disorder configurations. The results reported in this paper can be extended to more general models of homogeneous systems, but one should keep in mind the model of Eq. 6 as a prototypical example.

We end this section with a comment about discrete versus continuum models. Definitely, working with lattice models reduces tremendously the technical complexity of the problem. One crucial difference is that the algebra of observables does have a unit in the discrete case, and it doesn't in the continuum case. Also, when working with discrete models, we are dealing only with bounded observable but this will not be the case for continuum models. Hence, extending the theory to the continuum models will require more work at the technical level (note also that the results of Ref. [1], which is starting point for our work, were established for discrete models). However, even if we do extend the theory to continuum models, an explicit computer calculation for a real material will not be possible, simply because of the sheer complexity of the problem. This is the main reason the community working on topological materials focuses almost entirely on lattice models. We should mention that, at least for the periodic case, the Wannier function technique generates lattice models which exactly reproduce the bundle of the occupied states of the continuum models.

3. The non-commutative Brillouin torus (Bellissard [18])

3.1. The C^* -algebra of the covariant observables

If for perfectly periodic systems one is primarily concerned with the translationally invariant operators, for homogenous systems one is primarily concerned with the covariant families of operators. All relevant response functions and thermodynamic functions for a homogeneous model can be computed within the algebra generated by the covariant families of operators. This algebra can be given a C^* -algebra structure in the following way.

For $\{F_\omega\}_{\omega \in \Omega}$ a covariant family of operators, all the information is encoded in the matrix elements:

$$f_{\alpha,\beta}(\omega, \mathbf{n}) = \langle \mathbf{0}, \alpha | F_\omega | \mathbf{n}, \beta \rangle, \quad (12)$$

because all the other matrix elements can be obtained from $f_{\alpha,\beta}(\omega, \mathbf{n})$ using translations. Hence, instead of working with covariant families of operators, one can work with functions:

$$f : \Omega \times \mathbb{Z}^3 \rightarrow \mathcal{M}_{D \times D}, \quad (13)$$

where $\mathcal{M}_{D \times D}$ is the space of $D \times D$ complex matrices, and with the algebraic operations that follows from the operator algebra when written for the matrix elements of Eq. 12. These operations are:

$$\begin{aligned} (f + g)(\omega, \mathbf{n}) &= f(\omega, \mathbf{n}) + g(\omega, \mathbf{n}), \\ (f * g)(\omega, \mathbf{n}) &= \sum_{\mathbf{m} \in \mathbb{Z}^d} f(\omega, \mathbf{m}) g(\mathbf{t}_{\mathbf{m}}^{-1} \omega, \mathbf{n} - \mathbf{m}) e^{i\pi(\mathbf{n}, \hat{\mathbf{B}} \mathbf{m})}. \end{aligned} \quad (14)$$

For the beginning, one should consider only continuous functions f with compact support. The algebra generated by such elements will be called \mathcal{A}_0 . One can easily go back from the algebra \mathcal{A}_0 to the algebra of bounded operators using the operator representation π_ω given by the formula:

$$(\pi_\omega f) |\mathbf{n}\rangle = \sum_{\mathbf{m} \in \mathbb{Z}^d} e^{i\pi(\mathbf{m}, \hat{\mathbf{B}} \mathbf{n})} f(\mathbf{t}_{\mathbf{n}}^{-1} \omega, \mathbf{m} - \mathbf{n}) |\mathbf{m}\rangle, \quad (15)$$

where $|\mathbf{n}\rangle$ and $|\mathbf{m}\rangle$ should be viewed here as column matrices with entries $|\mathbf{m}, \alpha\rangle$ and $|\mathbf{m}, \beta\rangle$. The result of a representation $\{\pi_\omega f\}_{\omega \in \Omega}$ is automatically a covariant family of operators. For example, the Hamiltonian of Eq. 6 is generated by the element:

$$h_{\alpha,\beta}(\omega, \mathbf{n}) = t_{\mathbf{n}}^{\alpha\beta}(\omega). \quad (16)$$

If one introduces the norm on \mathcal{A}_0 :

$$\|f\| = \sup_{\omega \in \Omega} \|\pi_\omega f\|, \quad (17)$$

and the $*$ -operation:

$$f^*(\omega, \mathbf{n}) = f(\mathbf{t}_{\mathbf{n}}^{-1} \omega, -\mathbf{n})^\dagger, \quad (18)$$

then the completion of \mathcal{A}_0 under the norm of Eq. 17 becomes a C^* -algebra, which will be denoted by \mathcal{A} . The C^* -algebra structure allows one to develop the spectral problem and the functional calculus in a purely algebraic framework, without any reference to wavefunctions. The algebra \mathcal{A} is the smallest algebra that one can consider and still be able to compute the response functions of the homogeneous systems.

3.2. The non-commutative differential calculus

The classical differential manifolds are completely determined by the algebra of smooth functions defined over the manifold and by the differential calculus with these functions. In the non-commutative setting, one replaces the commutative algebra of smooth functions over the classic Brillouin torus with the non-commutative algebra \mathcal{A} , and then builds a non-commutative differential calculus over \mathcal{A} . For the latter, one replaces \mathbf{k} -integration over the classical Brillouin torus by a trace over \mathcal{A} :

$$\mathcal{T}(f) = \int_{\Omega} dP(\omega) \operatorname{tr}\{f(\omega, \mathbf{0})\}, \quad (19)$$

where tr is the trace over $\mathcal{M}_{D \times D}$ space, and the \mathbf{k} -derivations by linear automorphisms of algebra \mathcal{A} :

$$(\partial_j f)(\omega, \mathbf{n}) = \mathbf{n}_j f(\omega, \mathbf{n}), \quad j = 1, \dots, d. \quad (20)$$

Then the triplet $(\mathcal{A}, \mathcal{T}, \partial)$ determines a non-commutative manifold called the non-commutative Brillouin torus.

The trace \mathcal{T} has the standard properties:

- (i) Positivity: $\mathcal{T}(f * f^*) \geq 0$,
- (ii) Cyclic-city: $\mathcal{T}(f * g) = \mathcal{T}(g * f)$.

For the translational invariant case, \mathcal{T} reduces to the ordinary \mathbf{k} -integration, and for the general case with disorder and a magnetic field, the physical meaning of \mathcal{T} is that of the trace per unit volume:

$$\lim_{\Lambda \rightarrow \infty} \frac{1}{|\Lambda|} \operatorname{Tr}_{\Lambda}\{(\pi_{\omega} f)(\pi_{\omega} g) \dots\} = \mathcal{T}(f * g \dots), \quad (21)$$

where $\operatorname{Tr}_{\Lambda}$ means the trace over the quantum states inside the box Λ .

The derivations are not defined over the entire algebra \mathcal{A} but only over the sub-algebra of “differentiable” functions:

$$\mathcal{C}^1(\mathcal{A}) = \{f \in \mathcal{A}, \|\partial_j f\| < \infty, j = 1, 2, 3\}. \quad (22)$$

More general, the sub-algebra of N -th differentiable functions is defined as:

$$\mathcal{C}^N(\mathcal{A}) = \{f \in \mathcal{A}, \|\partial_1^{\alpha_1} \partial_2^{\alpha_2} \partial_3^{\alpha_3} f\| < \infty, \alpha_1 + \alpha_2 + \alpha_3 = N\}. \quad (23)$$

We will only work with $\mathcal{C}^{\infty}(\mathcal{A})$ elements, so from now all elements will be assumed to be part of this sub-algebra.

For convenience, we collect below a list of well-known rules of calculus ($j, k = 1, 2, 3$):

- (i) The derivations commute:

$$\partial_j \partial_k f = \partial_k \partial_j f. \quad (24)$$

- (ii) The derivations are *-derivations:

$$\partial_j(f^*) = (\partial_j f)^*. \quad (25)$$

- (iii) The derivations satisfy the Leibniz rule:

$$\partial_j(f * g) = (\partial_j f) * g + f * (\partial_j g). \quad (26)$$

- (iv) The operator representations of the derivations are:

$$\pi_{\omega}(\partial_j f) = -i[x_j, \pi_{\omega} f], \quad (27)$$

where $\mathbf{x} = (x_1, \dots, x_s)$ is the position operator.

(v) For f invertible in \mathcal{A} , then:

$$\partial_j f^{-1} = -f^{-1} * (\partial_j f) * f^{-1}. \quad (28)$$

(vi) If Φ and Φ' are any differentiable functions on the complex plane, then [21]:

$$\mathcal{T}(\Phi(f) * \partial_j \Phi'(f)) = 0. \quad (29)$$

In particular:

$$\mathcal{T}(\partial_j f) = 0. \quad (30)$$

If combined with the Leibnitz rule, one gets the partial integration rule:

$$\mathcal{T}(\partial_j f * g) = -\mathcal{T}(f * \partial_j g). \quad (31)$$

(vii) If Φ is an analytic function in a neighborhood of the unit circle and u is any unitary element ($u^* = u^{-1}$), then [22]:

$$\mathcal{T}(\Phi(u) \partial_j u) = a_{-1} \mathcal{T}(u^{-1} \partial_j u), \quad (32)$$

where a_{-1} is the coefficient of z^{-1} in the Laurent expansion $\Phi(z) = \sum_{n=-\infty}^{\infty} a_n z^n$.

4. The Ito derivative with respect to the magnetic field

As we already mentioned, the traditional linear response calculus with respect to the magnetic field can be very difficult. In the context of the magneto-electric response, this was already highlighted in Ref. [23]. In the non-commutative setting, however, this task is quite straightforward, in great part due to the pioneering work by Rammal and Bellissard [24]. According to this work, the derivative with respect to \mathbf{B} of the expected values can be computed as:

$$\partial_{B_j} \mathcal{T}(f) = \mathcal{T}(\delta_j f), \quad j = 1, 2, 3, \quad (33)$$

where δ_j 's are the Ito derivatives, which can be computed using the following rules of calculus (with j taken mod 3):

$$\delta_j \partial_k f = \partial_k \delta_j f, \quad (34)$$

$$\delta_j f^* = (\delta_j f)^*, \quad (35)$$

$$\begin{aligned} \delta_j(f * g) &= (\delta_j f) * g + f * (\delta_j g) \\ &+ \frac{i}{2} (\partial_{j+1} f * \partial_{j+2} g - \partial_{j+2} f * \partial_{j+1} g), \end{aligned} \quad (36)$$

$$\begin{aligned} \delta_j(f^{-1}) &= -f^{-1} * \delta_j f * f^{-1} \\ &+ \frac{i}{2} f^{-1} * [\partial_{j+1} f * f^{-1}, \partial_{j+2} f * f^{-1}]. \end{aligned} \quad (37)$$

These rules of calculus are enough to compute any response function involving the magnetic field. For example [1]:

Proposition 1. If p is an idempotent element ($p * p = p$), such as the projector onto the occupied states for our model, then:

$$p * (\delta_j p) * p = -\frac{i}{2} p * [\partial_{j+1} p, \partial_{j+2} p] * p \quad (38)$$

and

$$(1 - p) * (\delta_j p) * (1 - p) = \frac{i}{2} (1 - p) * [\partial_{j+1} p, \partial_{j+2} p] * (1 - p). \quad (39)$$

Proof. According to the rule in Eq. 36

$$\delta_j p = \delta_j(p * p) = (\delta_j p) * p + p * (\delta_j p) + \frac{i}{2}[\partial_{j+1} p, \partial_{j+2} p] \quad (40)$$

Then Eq. 38 follows after one multiplies the above equation with p on both left and right sides and collects all three resulting $p * \delta_j p * p$ terms. For the second identity, note that $(1 - p)$ is also an idempotent, so we can replace p by $(1 - p)$ in the first identity and use the fact that $\delta_j(1 - p) = -\delta_j p$. \square

The off-diagonal components of $\delta_j p$ do not have an expression that involves only p and its derivatives like in Eqs. 38 and 39. If p is the spectral projector of the Hamiltonian h onto the spectrum below the insulating gap:

$$p = \frac{i}{2\pi} \oint (h - z)^{-1} dz \quad (41)$$

where the contour encircles the occupied spectrum of h , then one will compute the off-diagonal part of $\delta_j p$ using the rule in Eq. 37 on $(h - z)^{-1}$. The result will depend on the particularities of h and will involve cross-gap matrix elements. In fact, in general, if we are dealing with a (analytic) function of the Hamiltonian, $\Phi(h)$, such as the Fermi-Dirac function at finite temperatures, then we can use the representation:

$$\Phi(h) = \frac{i}{2\pi} \oint \Phi(z)(h - z)^{-1} dz \quad (42)$$

and the rule in Eq. 37 on $(h - z)^{-1}$, to compute:

$$\begin{aligned} \delta_j \Phi(h) = \frac{i}{2\pi} \oint dz \Phi(z) \Big\{ & -(h - z)^{-1} * \delta_j h * (h - z)^{-1} \\ & + \frac{i}{2}(h - z)^{-1} [\partial_{j+1} h * (h - z)^{-1}, \partial_{j+2} h * (h - z)^{-1}] \Big\}. \end{aligned} \quad (43)$$

5. The isotropic magneto-electric response

The magneto-electric tensor is well defined only when referenced from a “standard” system. We will denote the reference system by h_0 and we will consider adiabatic time variations (*i.e.* infinitely slow) of the Hamiltonian $h(t)$ which originate at this h_0 and terminate at some h . The main assumption is that the Fermi level is always in a spectral gap during the adiabatic variations. At the technical level, one needs the assumption that $\partial_t h(t) = 0$ at the beginning and at the end of the adiabatic deformation. Now, let $p(t)$ be the projector onto the states below this spectral gap:

$$p(t) = (1 - \chi(h(t) - \epsilon_F))/2, \quad (44)$$

where $\chi(x) = \text{sign}(x)$. Under these conditions, Ref. [1] established the following result.

Proposition 2. The change of the electric polarization during an adiabatic variation can be computed via the formula:

$$\Delta P_j = i \int dt \mathcal{T}(p * [\partial_t p, \partial_j p]), \quad (45)$$

where the t -dependence of p is implied.

It is worthwhile to add a few more lines about the work of Ref. [1]. For the context of disordered systems under magnetic fields, these results are important because they lift the (super-) adiabatic theorem [25], including the full non-adiabatic error

estimation, to the non-commutative Brillouin torus. As such, one can foresee many additional applications which previously have been available only for the periodic case. Another notable point is that the formula for orbital magnetization derived in Ref. [1] does not require a spectral gap but only a mobility gap at the Fermi level. As such, there is hope that the theory of the magneto-electric response can be pushed into the regime of strong disorder. This is extremely important for the reasons to be discussed in the next sections.

Starting from the result stated in Proposition 1, we shall prove the following statement.

Proposition 3. In the above conditions, the variation in the isotropic magneto-electric response:

$$\Delta\alpha = \frac{1}{3} \sum_{j=1}^3 \partial\Delta P_j / \partial B_j \quad (46)$$

can be computed via :

$$\begin{aligned} \Delta\alpha = & \frac{1}{2} \int dt \varepsilon_{\alpha\beta\gamma\delta} \mathcal{T}(p * \partial_\alpha p * \partial_\beta p * \partial_\gamma p * \partial_\delta p) \\ & + \frac{i}{3} \mathcal{T}(\chi(h - \epsilon_F) * \partial_j p * \delta_j p) \Big|_{\text{initial}}^{\text{final}}, \end{aligned} \quad (47)$$

where summation over the repeating indices is assumed. The indices α, \dots, δ run through t and $j = 1, 2, 3$, and ε is the signature of the permutation.

Remark. The first (second) term on the right hand side of Eq. 47 is called topological (non-topological) for reasons to be discussed shortly. When the magnetic field and disorder are turned off, the topological part of $\Delta\alpha$ reduces to the formula derived in Refs. [4, 15, 16] in the context of perfectly periodic insulators. Showing a similar correspondence for the non-topological term is a more involved task and is deferred to future investigations.

Proof. Since:

$$\Delta\alpha = i \int dt \frac{1}{3} \sum_{j=1}^3 \mathcal{T}(\delta_j(p * [\partial_t p, \partial_j p])), \quad (48)$$

our main task is to compute:

$$\sum_{j=1}^3 \mathcal{T}(\delta_j(p * [\partial_t p, \partial_j p])). \quad (49)$$

Using the rule of calculus from Eq. 36, we have:

$$\begin{aligned} \delta_j(p * [\partial_t p, \partial_j p]) &= \delta_j p * [\partial_t p, \partial_j p] + p * \delta_j [\partial_t p, \partial_j p] \\ &+ \frac{i}{2} \partial_{j+1} p * \partial_{j+2} [\partial_t p, \partial_j p] - \frac{i}{2} \partial_{j+2} p * \partial_{j+1} [\partial_t p, \partial_j p]. \end{aligned} \quad (50)$$

The contributions to $\Delta\alpha$ from the last two terms above identically cancel out, as it can be easily seen by using an integration by parts (see Eq. 35). The Ito derivative of the commutator is:

$$\begin{aligned} \delta_j [\partial_t p, \partial_j p] &= [\delta_j \partial_t p, \partial_j p] + [\partial_t p, \delta_j \partial_j p] \\ &+ \frac{i}{2} ([\partial_{j+1} \partial_t p, \partial_{j+2} \partial_j p] - [\partial_{j+2} \partial_t p, \partial_{j+1} \partial_j p]). \end{aligned} \quad (51)$$

The contribution to the integrant in Eq. 48 from the first two terms of Eq. 51 is:

$$\begin{aligned} & \sum_{j=1}^3 \mathcal{T}(p * ([\delta_j \partial_t p, \partial_j p] + [\partial_t p, \delta_j \partial_j p])) \\ &= 2 \sum_{j=1}^3 \mathcal{T}([\partial_t p, \partial_j p] * \delta_j p) + \frac{d}{dt} \sum_{j=1}^3 \mathcal{T}(\chi(h - \epsilon_F) * \partial_j p * \delta_j p). \end{aligned} \quad (52)$$

Indeed:

$$\begin{aligned} & \mathcal{T}(p * [\delta_j \partial_t p, \partial_j p] + p * [\partial_t p, \delta_j \partial_j p]) \\ &= \mathcal{T}(p * \delta_j \partial_t p * \partial_j p - p * \partial_j p * \delta_j \partial_t p + p * \partial_t p * \delta_j \partial_j p - p * \delta_j \partial_j p * \partial_t p) \\ &= \mathcal{T}(\partial_j p * p * \delta_j \partial_t p - p * \partial_j p * \delta_j \partial_t p + p * \partial_t p * \delta_j \partial_j p - \partial_t p * p * \delta_j \partial_j p) \\ &= \mathcal{T}(\partial_t (\partial_j p * p * \delta_j p) - \partial_t \partial_j p * p * \delta_j p - \partial_j p * \partial_t p * \delta_j p \\ & \quad - \partial_t (p * \partial_j p * \delta_j p) + \partial_t p * \partial_j p * \delta_j p + p * \partial_t \partial_j p * \delta_j p \\ & \quad - \partial_j (p * \partial_t p) * \delta_j p + \partial_j (\partial_t p * p) * \delta_j p) \\ &= \partial_t \mathcal{T}((\partial_j p * p - p * \partial_j p) * \delta_j p) + 2 \mathcal{T}((\partial_t p * \partial_j p - \partial_j p * \partial_t p) * \delta_j p), \end{aligned} \quad (53)$$

and the affirmation follows because $\partial_j p * p = (1 - p) * \partial_j p$.

The contribution to the integrant in Eq. 48 from the last two terms of Eq. 51 is zero:

$$\sum_{j=1}^3 \mathcal{T}(p * ([\partial_{j+1} \partial_t p, \partial_{j+2} \partial_j p] - [\partial_{j+2} \partial_t p, \partial_{j+1} \partial_j p])) = 0. \quad (54)$$

Indeed, after an integration by parts, this contribution can be written as:

$$\begin{aligned} & \sum_{j=1}^3 \mathcal{T}(-\partial_{j+1} p * [\partial_t p, \partial_{j+2} \partial_j p] - p * [\partial_t, \partial_{j+1} \partial_{j+2} \partial_j p] \\ & \quad + \partial_{j+2} p * [\partial_t p, \partial_{j+1} \partial_j p] + p * [\partial_t, \partial_{j+2} \partial_{j+1} \partial_j p]) \\ &= \sum_{j=1}^3 \mathcal{T}(-\partial_{j+1} p * [\partial_t p, \partial_{j+2} \partial_j p] + \partial_{j+2} p * [\partial_t p, \partial_{j+1} \partial_j p]) \\ &= \mathcal{T}(-\partial_2 p * [\partial_t p, \partial_3 \partial_1 p] + \partial_3 p * [\partial_t p, \partial_2 \partial_1 p] \\ & \quad - \partial_3 p * [\partial_t p, \partial_1 \partial_2 p] + \partial_1 p * [\partial_t p, \partial_3 \partial_2 p] \\ & \quad - \partial_1 p * [\partial_t, \partial_2 \partial_3 p] + \partial_2 p * [\partial_t, \partial_1 \partial_3 p]), \end{aligned} \quad (55)$$

and in the final lines one can easily identify pairs of terms that identically cancel out.

So far the calculation stands as:

$$\begin{aligned} & \sum_{j=1}^3 \mathcal{T}(\delta_j (p * [\partial_t p, \partial_j p])) = 3 \sum_{j=1}^3 \mathcal{T}([\partial_t p, \partial_j p] * \delta_j p) \\ & \quad + \frac{d}{dt} \sum_{j=1}^3 \mathcal{T}(\chi(h - \epsilon_F) * \partial_j p * \delta_j p) \end{aligned} \quad (56)$$

Now:

$$\begin{aligned} & \mathcal{T}([\partial_t p, \partial_j p] * \delta_j p) \\ &= \mathcal{T}([\partial_t p, \partial_j p] * \delta_j p * p) + \mathcal{T}([\partial_t p, \partial_j p] * \delta_j p * (1 - p)) \\ &= \mathcal{T}(p * [\partial_t p, \partial_j p] * \delta_j p * p) + \mathcal{T}((1 - p) * [\partial_t p, \partial_j p] * \delta_j p * (1 - p)), \end{aligned} \quad (57)$$

where in the last line we used the cyclic-city of the trace and the idempotency of p and $1 - p$. Given the fact that p and $(1 - p)$ commute with $[\partial_t p, \partial_j p]$, we can continue as:

$$\dots = \mathcal{T}([\partial_t p, \partial_j p] * (p * \delta_j p * p + (1 - p) * \delta_j p * (1 - p))) \quad (58)$$

at which point we can use the rules in Eqs. 38 and 39 to continue as:

$$\begin{aligned} \dots &= \mathcal{T}([\partial_t p, \partial_j p] * (-\frac{i}{2} p * [\partial_{j+1} p, \partial_{j+2} p] * p \\ &\quad + \frac{i}{2} (1 - p) * [\partial_{j+1} p, \partial_{j+2} p] * (1 - p))) \\ &= \frac{1}{2i} \mathcal{T}((2p - 1) * [\partial_t p, \partial_j p] * [\partial_{j+1} p, \partial_{j+2} p]). \end{aligned} \quad (59)$$

It remains to demonstrate that:

$$\sum_{j=1}^3 \mathcal{T}([\partial_t p, \partial_j p] * [\partial_{j+1} p, \partial_{j+2} p]) = 0 \quad (60)$$

which follows after an explicit display of the terms involved in the sum:

$$\begin{aligned} &\sum_{j=1}^3 \mathcal{T}([\partial_t p, \partial_j p] * [\partial_{j+1} p, \partial_{j+2} p]) \\ &= \mathcal{T}(\partial_t p * \partial_1 p * \partial_2 p * \partial_3 p - \partial_t p * \partial_1 p * \partial_3 p * \partial_2 p \\ &\quad - \partial_1 p * \partial_t p * \partial_2 p * \partial_3 p + \partial_1 p * \partial_t p * \partial_3 p * \partial_2 p \\ &\quad + \partial_t p * \partial_2 p * \partial_3 p * \partial_1 p - \partial_t p * \partial_2 p * \partial_1 p * \partial_3 p \\ &\quad - \partial_2 p * \partial_t p * \partial_3 p * \partial_1 p + \partial_2 p * \partial_t p * \partial_1 p * \partial_3 p \\ &\quad + \partial_t p * \partial_3 p * \partial_1 p * \partial_2 p - \partial_t p * \partial_3 p * \partial_2 p * \partial_1 p \\ &\quad - \partial_3 p * \partial_t p * \partial_1 p * \partial_2 p + \partial_3 p * \partial_t p * \partial_2 p * \partial_1 p) \end{aligned} \quad (61)$$

and using the cyclic-city of the trace one can easily identify the pairs of terms that identically cancel out. At this point we arrived at:

$$\begin{aligned} \sum_{j=1}^3 \mathcal{T}(\delta_j(p * [\partial_t p, \partial_j p])) &= \frac{3}{i} \sum_{j=1}^3 \mathcal{T}(p * [\partial_t p, \partial_j p] * [\partial_{j+1} p, \partial_{j+2} p]) \\ &\quad + \frac{d}{dt} \sum_{j=1}^3 \mathcal{T}(\chi(h - \epsilon_F) * \partial_j p * \delta_j p), \end{aligned} \quad (62)$$

and the first term on the left side can be anti-symmetrized and brought to the desired form. This concludes the proof. \square

6. Quantization and Invariance

For the purpose of this discussion, it is more convenient to view the adiabatic time evolution from an initial to a final configuration as an adiabatic interpolation between two points in the parameter space of the model. The adiabatic interpolations will be represented as curves in this parameter space and the time-integral can be viewed as an integral along such curve. We automatically assume that the spectral gap at the Fermi level remains open during the adiabatic interpolations. We will use the notations $\gamma + \gamma'$ to denote the interpolation path obtained by joining the interpolation

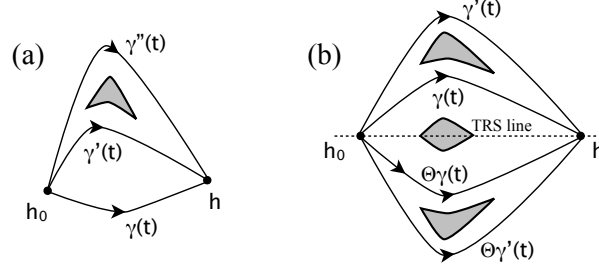


Figure 1. (a) Various adiabatic interpolations between the initial (h_0) and final (h) systems. (b) Here h_0 and h are time-reversal symmetric (TRS), but they are adiabatically interpolated via paths that are not necessarily time-reversal symmetric ($\Theta\gamma \neq \gamma$). The shaded regions represents regions in the model's parameter space where the spectral gap closes. The TRS line is invariant to the time-reversal operation.

paths γ and γ' , $-\gamma$ to denote the interpolation path γ when taken in reverse, and $\Delta\alpha[\gamma]$ to indicate that the variation in the magneto-electric response depends on the interpolation path γ .

The first question we want to address is what happens with the variation of the isotropic magneto-electric response when one considers different interpolation paths, $\gamma(t)$ and $\gamma'(t)$, between the reference and the final systems. Since:

$$\Delta\alpha[\gamma] - \Delta\alpha[\gamma'] = \Delta\alpha[\gamma - \gamma'], \quad (63)$$

we can write:

$$\Delta\alpha[\gamma] - \Delta\alpha[\gamma'] = \frac{1}{2} \oint_{\gamma - \gamma'} dt \varepsilon_{\alpha\beta\gamma\delta} \mathcal{T}(p * \partial_\alpha p * \partial_\beta p * \partial_\gamma p * \partial_\delta p), \quad (64)$$

and the important observation here is that the t -integral is now taken over a closed loop. One can then recognize on the right-hand side the expression of the second Chern number C_2 over the manifold $(\gamma - \gamma') \times \text{Brillouin torus}$, written in the real-space rather than in the k -space. Because of the finite-gap assumption, the second Chern number is quantized, a fact that follows from a general argument by Connes [26], which says that the pairing between the projectors and cyclic cocycles is a homotopy invariant. Hence the quantization occurring in the clean limit persists when the disorder is turned on, at least as long as the spectral gap remains open.

Therefore, the answer to our first question is that the difference in $\Delta\alpha[\gamma]$, when computed along different adiabatic evolutions, is always an integer number (we recall that our units were taken such that $e = h = 1$). We illustrate in Fig. 1(a) various possible situations. For example, if the interpolating paths $\gamma(t)$ and $\gamma'(t)$ can be deformed into each other without closing the spectral gap, then C_2 is necessarily zero and $\Delta\alpha[\gamma]$ is identical for the two interpolation paths. But, if the interpolating paths cannot be deformed into each other without closing the spectral gap, such as it happens for $\gamma(t)$ and $\gamma''(t)$ in Fig. 1(a), then C_2 may take non-trivial integer values so $\Delta\alpha[\gamma]$ and $\Delta\alpha[\gamma'']$ cannot be expected to be identical, but we can say with absolute confidence that the difference between the two will be an integer number.

We now turn our attention to adiabatic interpolations between time-reversal invariant systems, in which case we need to turn the magnetic field off. Symmetry

considerations can be used to show that only the topological part of $\Delta\alpha$ survives for such systems [15, 16]. The time-reversal operation is given by the map:

$$\Theta : \mathcal{A}|_{B=0} \rightarrow \mathcal{A}|_{B=0}, \quad (\Theta f)(\omega, \mathbf{n}) = e^{i\pi s_y} \overline{f(\omega, \mathbf{n})} e^{-i\pi s_y}, \quad (65)$$

where s_y is a matrix from $\mathcal{M}_{D \times D}$ representing the y -component of the spin operator, and the line at the top represents complex conjugation. Some immediate properties of time-reversal map Θ are:

$$\Theta(f * g) = (\Theta f) * (\Theta g), \quad (66)$$

and

$$\partial_j \Theta f = -\Theta(\partial_j f). \quad (67)$$

Now, let h_0 and h be two time-reversal invariant Hamiltonians, $\Theta h_0 = h_0$ and $\Theta h = h$. The key observation is that, if $\gamma(t)$ is an adiabatic interpolation between h_0 and h , then $\Theta\gamma$ is also an adiabatic interpolation between h_0 and h (with a strictly positive spectral gap), and furthermore:

$$\Delta\alpha[\gamma] = \Delta\alpha[-\Theta\gamma]. \quad (68)$$

This can be seen from the explicit formula of Eq. 47 and the properties of the time-reversal map stated in Eqs. 66 and 67. We already established that:

$$\Delta\alpha[\gamma] - \Delta\alpha[\Theta\gamma] = C_2[\gamma - \Theta\gamma], \quad (69)$$

and this together with Eq. 68 leads to an important conclusion for this section:

$$\Delta\alpha[\gamma] = \frac{1}{2} C_2[\gamma - \Theta\gamma]. \quad (70)$$

This is a generalization of the Qi-Hughes-Zhang argument [4] to the case of disordered insulators.

In Fig. 1(b) we illustrate several possible situations. If γ can be deformed to a time-reversal invariant path, $\Theta\gamma = \gamma$, then $C_2 = 0$ and $\Delta\alpha[\gamma]$ is necessarily zero. But if h_0 and h cannot be connected with a time-reversal invariant path without closing the spectral gap at the Fermi level, then γ must be taken as in Fig. 1(b) so that to avoid the closing of the spectral gap. In this case, $C_2[\gamma - \Theta\gamma]$ can take non-trivial integer values and consequently $\Delta\alpha[\gamma]$ can take non-trivial values, which are either integer or half-integer numbers. If we take a different interpolation path, such as γ' in Fig. 1(b), then $\Delta\alpha[\gamma']$ may change but only by an integer number, as it was already established above. As such, for all possible adiabatic interpolation paths, $\Delta\alpha[\gamma]$ is either an integer number or a half-integer number. In other words, the integer or half-integer character of $\Delta\alpha[\gamma]$ is topologically stable.

7. Conclusions

The results of the last section can be further developed into a \mathbb{Z}_2 classification of the 3-dimensional disordered time-reversal invariant insulators. For this, one can take the reference systems to be the insulators with zero hopping amplitudes between different lattice sites. These systems can be called “trivial” since they do not display any magneto-electric response as there are no Peierls factors when a magnetic field is turned on. As such, we can postulate $\alpha = 0$ for the trivial systems. Then, the time-reversal invariant insulators fall into two classes (here we assume that all time-reversal symmetric insulators can be adiabatically connected to a representative from

the trivial class): one for which $\alpha = \text{integer}$ and one for which $\alpha = \text{half-integer}$, and we can say with absolute confidence that an insulator from the first class and an insulator from the second class cannot be connected by a time-reversal invariant interpolation path without closing the insulating gap. An insulator from a given symmetry class is generally called topological [6] if it cannot be adiabatically connected with its atomic limit (obtained by turning off the hopping terms between distinct lattice sites) via an interpolation path that respects the symmetry of the class and without closing the insulating gap. If we follow this definition, then the class of 3-dimensional disordered time-reversal invariant insulators with $\alpha = \text{half-integer}$ contains only topological insulators. Of course, our analysis does not rule out additional, more refined topological sub-classifications of either $\alpha = \text{half}$ or $\alpha = \text{half-integer}$ classes.

A \mathbb{Z}_2 classification of the disordered inversion symmetric insulators is also possible via similar arguments. This is the case because the inversion operation satisfies the properties written in Eqs. 66 and 67, which are at the core of the \mathbb{Z}_2 classification of time-reversal insulators. However, for inversion symmetric insulators, it is explicitly known [6] that the \mathbb{Z}_2 classification is not the end of the story as it can be further refined.

We would like to point out that the computation of Section 5 remains valid even in the regime of strong disorder where the insulating gap is closed and is replaced by a mobility gap (provided we accept the formula Eq. 45 for the electric polarization). If a non-commutative theory of the second Chern number paralleling that of the first Chern number [21] can be developed, then the magneto-electric response can be used to topologically classify the time-reversal and inversion symmetric insulators in the strong disorder regime, which is one of the long-standing open problems in the field.

The non-commutative formula Eq. 47 of the magneto-electric response can be efficiently evaluated on a computer by using the methods developed in Ref. [27]. As such, we now have a practical tool to explore, at least numerically, the physics of strongly disordered 3-dimensional topological insulators. The previous numerical attempts [28] to compute the 3-dimensional \mathbb{Z}_2 invariants based on twisted boundary methods were limited to small systems due to the extreme complexity of such methods, even though these calculations involved a gauge invariant, highly efficient formulation of the \mathbb{Z}_2 invariant [29]. From our experience with other non-commutative formulas (such as those for the first Chern number and for the Kubo formula [30, 27, 31]), we are quite optimistic that the non-commutative formula for magneto-electric polarization will enabled computations for much larger volumes where finer and more accurate results can be obtained.

Acknowledgments

We thank David Vanderbilt for extremely useful discussions. We also thank the reviewers for extremely useful comments. This work was supported by the U.S. NSF grants DMS-1066045 and DMR-1056168.

References

- [1] Schulz-Baldes H and Teufel S 2012 arXiv:1201.4812v1
- [2] Fiebig M 2005 *J. Phys. D: Appl. Phys.* **38** R123–R152
- [3] Spaldin N A and Fiebig M 2005 *Science* **309** 391–392
- [4] Qi X L, Hughes T L and Zhang S C 2008 *Phys. Rev. B* **78** 195424
- [5] Turner A M, Zhang Y, Mong R S K and Vishwanath A 2012 *Phys. Rev. B* **85** 165120
- [6] Hughes T, Prodan E and Bernevig B A 2011 *Phys. Rev. B* **83** 245132

- [7] Hasan M Z and Kane C L 2010 *Rev. Mod. Phys.* **82** 3045–3067
- [8] Qi X L and Zhang S C 2011 *Rev. Mod. Phys.* **83** 1057–1110
- [9] Wang Z, Qi X and Zhang S C 2010 *New J. Phys.* **12** 065007
- [10] Fu L, Kane C L and Mele E J 2007 *Phys. Rev. Lett.* **98** 106803
- [11] Moore J E and Balents L 2007 *Phys. Rev. B* **75** 121306
- [12] Roy R 2009 *Phys. Rev. B* **79** 195322
- [13] Li R, Wang J, Qi X L and Zhang S C 2010 *Nature Phys.* **6** 284–288
- [14] Essin A M, Moore J E and Vanderbilt D 2009 *Phys. Rev. Lett.* **102** 146805
- [15] Essin A M, Turner A M, Moore J E and Vanderbilt D 2010 *Phys. Rev. B* **81** 205104
- [16] Malashevich A, Souza I, Coh S and Vanderbilt D 2010 *New J. Phys.* **12** 053032
- [17] Connes A 1994 *Noncommutative Geometry* (San Diego, CA: Academic Press)
- [18] Bellissard J 1986 *Lecture Notes in Physics* vol 257 ed Dorlas T, Hugenholtz M and Winnink M (Springer-Verlag) pp 99–156
- [19] Liu C X, Qi X L, Zhang H, Dai X, Fang Z and Zhang S C 2010 *Phys. Rev. B* **82** 045122
- [20] Panati G, Spohn H and Teufel S 2003 *Commun. Math. Phys.* **242** 547–578
- [21] Bellissard J, van Elst A and Schulz-Baldes H 1994 *J. Math. Phys.* **35** 5373–5451
- [22] Prodan E 2009 *J. Phys. A: Math. Theor.* **42** 065207
- [23] Chen K T and Lee P A 2011 *Phys. Rev. B* **84** 205137
- [24] Rammal R and Bellissard J 1990 *J. Phys. France* **51** 1803–1830
- [25] Nenciu G 1993 *Comm. Math. Phys.* **152** 479–496
- [26] Connes A 1985 *Publications Mathématiques de l’I.H.E.S.* **62** 257–360
- [27] Prodan E 2012 *Appl. Math. Res. Express* **2012** abs017
- [28] Leung B and Prodan E 2012 *Phys. Rev. B* **85** 205136
- [29] Prodan E 2011 *Physical Rev. B* **83** 235115
- [30] Prodan E 2011 *J. Phys. A: Math. Theor.* **44** 113001
- [31] Xue Y and Prodan E 2012 *Phys. Rev. B* **86** 155445